Estimating the Crowd Size of a Rally by Crowdsourcing-Geocomputation

T. Edwin Chow [a], *

[a] *Texas State University, chow@txstate.edu,*

* Corresponding author

Abstract:

Dynamic population estimation – counting people in a special event (e.g. rally, sport events, evacuation) has been challenging because a large crowd is difficult to acquire an accurate count manually as people can join and leave the crowd at any time and place. In the context of a rally where people are moving on the streets over an extended period, it is often difficult to answer some questions with regards to confine the phenomenon, including but not limited to: When and where does the rally start/end precisely? Who are the attendees (e.g. activists, spectators, organizers, police, opposing crowd(s), etc.)? How does the confining environment affect stochastic human behaviours, interactions and movements over time and space? What is the count (e.g. peak attendance, total count)?

Nevertheless, estimating the attendance of these dynamic events could be emotionally and politically charged. Due to these practical uncertainties of dynamic population estimation, the attendance of annual July 1st Rally in Hong Kong reported by the police and organizers could be very different as it often became a bluffing game to promote hidden political agenda. For example, the police estimated a peak attendance of 9,800 in the 2017 July 1st Rally, whereas the organizers reported five times the attendance at about 50,000. Independent researchers from the University of Hong Kong estimated the total attendance of 14,170 (Yip, 2017) and 29,000 (HKUPOP, 2017) respectively. With limited field observations, it is hard to examine the accuracy and confident level of these reported counts.

This research aims to estimate the total attendance of 2017 July 1st Rally in Hong Kong and examine the counts reported by various stakeholders. Based on many stories reported by the public and social media, the timeline of the rally event was reconstructed to trace the observed rally time of the head and tail crowds participated in the rally event. Important geospatial features, including the street network, approved protest areas and entry/exit gateways along the main rally route, were reconstructed in the Geographic Information System (GIS). This study adopted a crowdsourcing-geocomputation approach to simulate how a dynamic crowd would have navigated in such as rally event (Chow, 2019). Using a mobile application that tracks individual trajectory, volunteers were recruited to contribute valuable in-situ data of dynamic human movements and behaviours attending the rally event. These data were used to formulate and calibrate the parameters of a computational cartographic model where each rally attendees were represented as a moving agent (x, y, t) confined in a micro spatial environment where the protesters marched from Victoria Park to the Government Headquarter in Hong Kong. Hence, individual GPS (Global Positioning System) trajectory during the rally was collected and converted into GIS data format for further analysis. More details about the data collection and processing can be found on the project website (https://chowte.wixsite.com/dynamicpop). By leveraging a large number of observations volunteered by crowdsourcing, this study attempted to answer the research question: *What might be the simulated crowd size that reasonable range of model parameters can be converged?*

Using this crowdsoucing-geocomputation model, a sensitivity analysis was conducted to simulate varying model parameters, including maximum walking speed, maximum crowd density, early departure and late entry rates. Based on the reported count reported by the police, organizers and independent researchers, various crowd sizes were simulated to be compared against the observed rally time of 209 minutes (i.e. about 3.5 hours) from start to finish. Using the crowdsourced data for calibration, most rally models simulated an arrival time of head crowd between 106 – 108 minutes, which was very close to the observed rally time of head crowd of 107 minutes. In this study, crowd sizes were considered to be acceptable based on a 95% confidence interval of arrival time of tail crowd (i.e. 174 – 192 minutes) and total rally time (i.e. 199 – 219 minutes). For example, a crowd size of 9,800 people was simulated matching the total rally time to examine the credibility of calibrated model parameters (Figure 1).
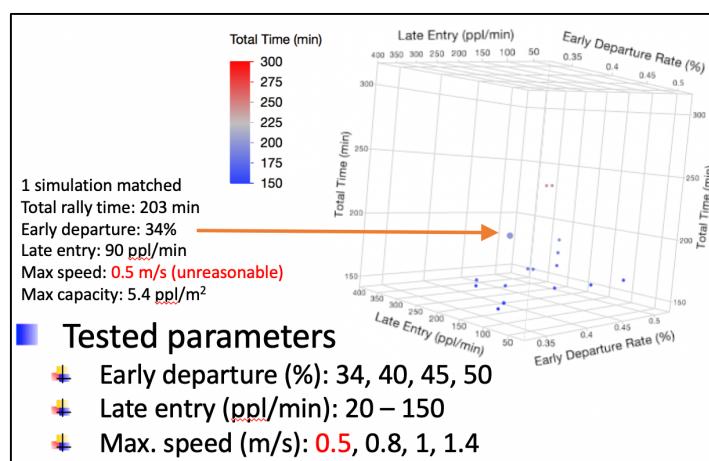
Figure 1. A simulation of 9,800 people reported by the police matching the observed total rally time and its parameters.

Within the tested range of calibrated model parameters, the results indicated that it was possible to tweak the model parameters of varying crowd size to match the observed rally time (Table 1). Despite the simulated rally time of some accepted models were within ±5% of the observed rally time, the parameters used to simulate such a model were not necessary reasonable in reality. The simulated count of 9,800, for example, would require a cap of maximum walking speed of 0.5 m/s, which seemed to be unreasonably slow under normal circumstance and incompatible with crowdsourced data (Figure 1). Given the observed rally time, it was found that the crowd sizes of 14,000 – 29,000 could be simulated with reasonable model parameters, whereas the crowd sizes of 9,800 and 50,000 would yield unreasonable model parameters. Taking the median within the range of 14,000 – 29,000, this study also found that a crowd size of 21,000 could yield eight matching simulations with varying reasonable model parameters that may be better simulate the actual rally attendance.

Stakeholders	Reported Count	Matches Simulation
Police	9,800	1
Yip	14,170	2
HKUPOP	29,000	1
Civil Human Right Front	50,000	1

Table 1. Matching simulations of varying crowd sizes based on the observed rally time.

This paper provided empirical evidences to examined the credibility of various crowd sizes of the 2017 July 1st Rally in Hong Kong reported by the stakeholders. The research also presents a transparent, repeatable and verifiable approach to explore, quantify and simulate human movements in a rally event, such as the early departure and late arrival, to better understand dynamic crowd behaviours and interactions.

References

Chow, T.E. (2019). A crowdsourcing-geocomputational framework of mobile crowd simulation and estimation, *Cartography and Geographic Information Science*, 46(1): 2-20. DOI: 10.1080/15230406.2018.1524314.

HKUPOP, (2017) July 1 Rally. URL: https://www.hkupop.hku.hk/english/features/july1/headcount/2017/index.html, Last visited on January 19, 2019.

Yip, S.F.P. (2017) personal communication, July 5, 2017.